

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2000-29585

(P 2 0 0 0 - 2 9 5 8 5 A)

(43) 公開日 平成12年1月28日(2000.1.28)

(51) Int. Cl. <sup>7</sup>	識別記号	F I	テーマコード (参考)
G06F 3/00	601	G06F 3/00 601	2C061
B41J 29/17		G03G 21/00 370	2H027
G03G 21/00	370	G06F 3/12	R 5B021
G06F 3/12		G10L 3/00 551	Z 5C062
G10L 15/00		H04N 1/00	C 5D015

審査請求 未請求 請求項の数 6 O L (全16頁) 最終頁に続く

(21) 出願番号 特願平10-193069

(22) 出願日 平成10年7月8日(1998.7.8)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 青柳 剛

東京都大田区下丸子3丁目30番2号 キヤ  
ノン株式会社内

(74) 代理人 100066061

弁理士 丹羽 宏之 (外1名)

最終頁に続く

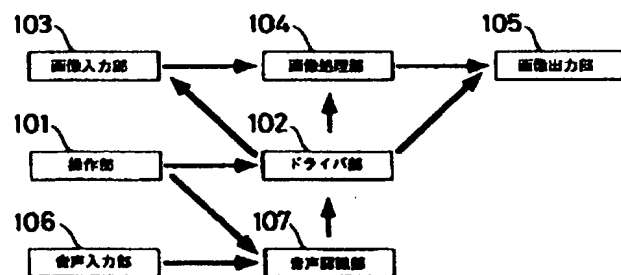
(54) 【発明の名称】 音声コマンド認識画像処理装置

(57) 【要約】

【課題】 音声入力コマンドを認識して作動する複写機及びファクシミリやプリンタ等の複合機において、不特定多数の人が使用する場合、予め登録する音声辞書データ用のメモリ等の記憶媒体の容量が過大となることを緩和にするための手段を提供する。

【解決手段】 このため、辞書データに参照データを登録する際は、実際に操作部101のボタン等を押して登録し、次回同じコマンドを使用する時には、音声のみで操作できるようにし、また、登録を行う際にもパターンマッチングを行い、既登録済みコマンドと同一または類似コマンドや類似音声が入力された場合には、その使用者に警告すると共に、また、類似度の高い辞書データは、指紋や掌紋等の他の判断要素でより正確に認識するように構成した。

各実施例の複写機の概要構成ブロック図



## 【特許請求の範囲】

【請求項 1】 スイッチやボタン、テンキー、液晶タッチパネル等により、コピー操作やファクシミリの送信等のコマンドの入力を行う操作手段と、

CPU等の処理装置を使用し、前記操作手段からのコマンドにより画像処理の設定やプリント枚数等の設定を行い、画像入力手段、画像処理手段、画像出力手段等の動作の制御を行うドライバ手段と、

CCD等を使用し、前記ドライバ手段からの制御により、原稿の画像データをデジタルデータに変換し、画像 10 処理部へ出力する画像入力手段と、

ASIC等のハード回路や、前記CPU等を使用したソフト的な処理等により、前記画像入力手段から入力されるデジタル画像データを、前記ドライバ手段からの画像処理の設定に応じて処理を行い、前記画像出力手段に出力を行う画像処理手段と、

前記画像処理手段により処理された画像データを、レーザプリンタやインクジェットプリンタ等のプリンタによって出力を行う画像出力手段であって、この部分はブラウン管モニタや液晶モニタ、または前記操作手段部に使用されている前記液晶タッチパネルに表示を行ったり、画像データとしてネットワーク上への出力、モデムを介してファクシミリ送信等を行うものも含む前記画像出力手段と、マイク等を使用し、音声の入力を行い、前記デジタルデータとして音声認識手段に送る音声入力手段と、前記音声入力手段からの音声データを入力し、辞書データへの登録や音声のパターンマッチング、コマンドへの変換等を行う音声認識手段と、

より構成され、前記音声によるコマンドの認識を行い動作することを特徴とする音声コマンド認識画像処理装置。 30

【請求項 2】 前記音声認識手段は、前記音声入力手段からのデジタル音声データを入力する音声データ入力手段と、前記操作手段から入力されたコマンド情報の入力を行う操作部情報入力手段と、前記操作手段部情報と前記音声データとを対応させて、メモリやハードディスク等の記憶媒体に記憶しておく辞書データ蓄積手段と、前記音声データ入力手段からのデータと、前記辞書データ蓄積手段からの音声辞書データとのパターンマッチングを行い、各コマンドに対する類似度を出力するパターンマッチング手段と、前記辞書情報の登録や通常コピー動作の情報を前記ドライバ手段から入力を行うドライバ情報入力手段と、前記パターンマッチング手段からの情報と、前記コマンド情報入力手段の情報とを受けて、通常コピーの場合は、入力された前記音声コマンドがいずれのコマンドに該当するかを判断し、コマンド出力手段に送るコマンド処理手段と、

前記コマンド処理手段からのコマンド情報を、通常コピーの場合はドライバ手段や操作手段のモード表示部分へ送るコマンド出力手段と、

より成ることを特徴とする請求項 1 記載の音声コマンド認識画像処理装置。

【請求項 3】 前記音声コマンド認識画像処理装置は、前記辞書データ蓄積手段にデータを登録する際に、実際に操作しながら前記音声データの登録を行うことを特徴とする請求項 1、2 のいずれか記載の音声コマンド認識画像処理装置。

【請求項 4】 前記音声コマンド認識画像処理装置は、前記音声データを登録する際に、前記辞書データ蓄積手段にデータに既登録のデータと同一データを登録しようとした場合、その旨を使用者に報知することを特徴とする請求項 1、2、3 のいずれか記載の音声コマンド認識画像処理装置。

【請求項 5】 前記音声コマンド認識画像処理装置は、前記音声により設定を行う際に、未登録のコマンドを前記音声により設定しようとした場合、その未登録の旨を使用者に報知することを特徴とする請求項 1 ないし 4 のいずれか記載の音声コマンド認識画像処理装置。

【請求項 6】 前記音声コマンド認識画像処理装置は、前記音声により設定を行う際に、その音声入力されたコマンドが複数のコマンドと類似していた場合、それらのコマンドを報知し、使用者に選択を促すことを特徴とする請求項 1 ないし 5 のいずれか記載の音声コマンド認識画像処理装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、複写機及びファクシミリやプリンタ等の複合機、特にコマンドを入力するユーザインタフェースに音声を用いているこの種の複合機械に関するものである。

## 【0002】

【従来の技術】近年、パーソナルコンピュータやカーナビゲーション等に見られる様に、音声認識を利用したコマンド入力を行う機器が普及しつつあるが、従来の複写機及びファクシミリやプリンタ等の複合機には、以上のような機能を有するものはなかった。

【0003】この種の音声認識を行うには、声紋や音声の区切り、音の高低等の“パターンマッチング”を行っている。このパターンマッチングとは、予めメモリ等の記憶媒体に使用するコマンドと、それを発声した時の音声データとが 1 対 1 に対応して、辞書として登録されており、入力された音声データが登録されている各コマンドと、どの程度一致しているかを類似度として算出し、その値の最大のものをコマンドとして選択し、実行するという類のものである。

## 【0004】

50 【発明が解決しようとする課題】しかしながら、前記複

写機及びファクシミリやプリンタ等の複合機において、前記のような音声認識によるコマンド入力を行うに当たっては、以下のような問題点がある。

【0005】すなわち、音声認識でコマンドの入力を行う際には、予め使用する全てのコマンドを辞書に登録しておく必要がある。現在のコマンドの登録方法は、使用するコマンドの全てを1度読み乍ら、各キー操作と対応させて登録していかななくてはならなかった。

【0006】従来のように、ある程度使用者が限定されているカーナビゲーションやパーソナルコンピュータ等では、登録するデータは概して個人的のものであり、1人の音声データを入力するだけで良かった。

【0007】しかしながら、複写機のように不特定多数の人が使用する機械の場合には、使用する可能性のある多くの人の音声辞書データを予め登録しておく必要があり、そのデータを記憶するために使用するメモリ等の記憶媒体の容量もそれなりに大きなものが必要となる。

【0008】また、登録しておくコマンドも、人それぞれ使用頻度も異なり、人によっては結果的には全く使用しないコマンドがあっても、全てを登録しておく必要があり、それによっても結果的には、不要な記憶媒体の容量を使用することとなってしまう。

【0009】本発明は、以上のような局面にかんがみてなされたもので、これらの問題点を解決することのできる音声コマンド認識画像処理装置の提供を目的としている。

#### 【0010】

【課題を解決するための手段】このため、本発明においては、以下の各項(1)～(6)のいずれかに示すような音声コマンド認識画像処理装置を提供することにより、前記目的を達成しようとするものである。

【0011】(1) スイッチやボタン、テンキー、液晶タッチパネル等により、コピー操作やファクシミリの送信等のコマンドの入力を行う操作手段と、CPU等の処理装置を使用し、前記操作手段からのコマンドにより画像処理の設定やプリント枚数等の設定を行い、画像入力手段、画像処理手段、画像出力手段等の動作の制御を行うドライバ手段と、CCD等を使用し、前記ドライバ手段からの制御により、原稿の画像データをデジタルデータに変換し、画像処理部へ出力する画像入力手段と、ASIC等のハード回路や、前記CPU等を使用したソフト的な処理等により、前記画像入力手段から入力されるデジタル画像データを、前記ドライバ手段からの画像処理の設定に応じて処理を行い、前記画像出力手段に出力を行う画像処理手段と、前記画像処理手段により処理がされた画像データを、レーザプリンタやインクジェットプリンタ等のプリンタによって出力を行う画像出力手段であって、この部分はブラウン管モニタや液晶モニタ、または前記操作手段部に使用されている前記液晶タッチパネルに表示を行ったり、画像データとしてネットワー

ク上への出力、モデムを介してファクシミリ送信等を行うものも含む前記画像出力手段と、マイク等を使用し、音声の入力を行い、前記デジタルデータとして音声認識手段に送る音声入力手段と、前記音声入力手段からの音声データを入力し、辞書データへの登録や音声のパターンマッチング、コマンドへの変換等を行う音声認識手段と、より構成され、前記音声によるコマンドの認識を行い動作することを特徴とする音声コマンド認識画像処理装置。

10 【0012】(2) 前記音声認識手段は、前記音声入力手段からのデジタル音声データを入力する音声データ入力手段と、前記操作手段から入力されたコマンド情報の入力を行う操作部情報入力手段と、前記操作手段部情報と前記音声データとを対応させて、メモリやハードディスク等の記憶媒体に記憶しておく辞書データ蓄積手段と、前記音声データ入力手段からのデータと、前記辞書データ蓄積手段からの音声辞書データとのパターンマッチングを行い、各コマンドに対する類似度を出力するパターンマッチング手段と、前記辞書情報の登録や通常コピー動作の情報を前記ドライバ手段から入力を行うドライバ情報入力手段と、前記パターンマッチング手段からの情報と、前記コマンド情報入力手段の情報とを受け、通常コピーの場合は、入力された前記音声コマンドがいずれのコマンドに該当するかを判断し、コマンド出力手段に送るコマンド処理手段と、前記コマンド処理手段からのコマンド情報を、通常コピーの場合はドライバ手段や操作手段のモード表示部分へ送るコマンド出力手段と、より成ることを特徴とする前項(1)記載の音声コマンド認識画像処理装置。

30 【0013】(3) 前記音声コマンド認識画像処理装置は、前記辞書データ蓄積手段にデータを登録する際に、実際に操作しながら前記音声データの登録を行うことを特徴とする前項(1)、(2)のいずれか記載の音声コマンド認識画像処理装置。

40 【0014】(4) 前記音声コマンド認識画像処理装置は、前記音声データを登録する際に、前記辞書データ蓄積手段にデータに既登録のデータと同一データを登録しようとした場合、その旨を使用者に報知することを特徴とする前項(1)、(2)、(3)のいずれか記載の音声コマンド認識画像処理装置。

【0015】(5) 前記音声コマンド認識画像処理装置は、前記音声により設定を行う際に、未登録のコマンドを前記音声により設定しようとした場合、その未登録の旨を使用者に報知することを特徴とする前項(1)ないし(4)のいずれか記載の音声コマンド認識画像処理装置。

50 【0016】(6) 前記音声コマンド認識画像処理装置は、前記音声により設定を行う際に、その音声入力されたコマンドが複数のコマンドと類似していた場合、それらのコマンドを報知し、使用者に選択を促すことを特徴

とする前項(1)ないし(5)のいずれか記載の音声コマンド認識画像処理装置。

【0017】

【作用】以上のような本発明構成により、前記問題点を解決した、新しい音声によるコマンド入力を行うことができる。すなわち、

1) 辞書データに参照するデータを登録する際は、実際に操作部のボタン等を押して登録することによって登録を行い、次回同じコマンドを使用する時には、音声のみで操作ができるようにすることによって、予め時間を割いて各使用者が全てのコマンドの登録を行う必要がなくなると共に、その機械を使用する各人が頻繁に使用するコマンドのみを登録することができる。これによって辞書に使用する記憶媒体の容量の削減を実現し得る。

【0018】2) また、データ登録を行う際にも、パターンマッチングを行い、既登録のコマンドと同じコマンドを入力した時や、類似したコマンド、類似した音声が入力された場合には、表示や警告音等によって使用者に報知することにより、登録しようとしたコマンドが、正  
20 当に入力されるべきものであるか否かを確認することができる。

【0019】3) また、音声による使用者の認識を行う際に、登録してある辞書データに、ある程度類似度の高いものが複数あった場合には、他の判定要素、例えば指紋や掌紋などによって、より正確に認識を行うようにすることにより、音声のみで使用者の認識を行うより、さらに正確に個人の識別を行うことが可能となる。

【0020】

【発明の実施の形態】以下に、本発明の実施の形態を、複数の実施例に基づき、それぞれ図面を参照して詳細に  
30 説明する。

【0021】

【実施例】(第1の実施例)本発明に関する画像処理装置の具体的な第1の実施例を、複写機の事例を用いて説明する。図1に、本第1実施例の複写機の構成ブロック図を示す(図1は、後述する第2～第5実施例にも共用するものとする)。

【0022】101は操作部で、図8にその具体的な平面図例を示すように、各スイッチやボタン、テンキー、  
40 液晶タッチパネル等によりコピー操作やファクシミリの送信等のコマンドの入力等を行う。図8における個々の構成要素の説明は、それぞれ図示註記に渡り、重複説明は省略する。

【0023】102はドライバ部で、CPU等の処理装置を使用し、操作部101からのコマンドによって、画像処理の設定やプリント枚数等の設定を行い、スキャナ部103、画像処理部104、プリンタ部105等の動作の制御を行う。

【0024】103は画像入力部で、CCD等を使用し、ドライバ部102からの制御によって、原稿の画像  
50

データをデジタルデータに変換し、画像処理部104へと出力する。

【0025】104は画像処理部で、ASIC等のハード回路や、CPU等を使用したソフト的な処理等によって、画像入力部103から入力されてくるデジタル画像データを、ドライバ部102からの画像処理の設定に応じて処理を行い、画像出力部105に出力を行う。

【0026】105は画像出力部で、画像処理部104によって処理が行われた画像データを、レーザプリンタやインクジェットプリンタ等のプリンタによって出力を行ったり、ブラウン管モニタや液晶モニタ、または操作部101に使用されているタッチパネル等に出力を行う。また、画像データとしてネットワーク上に出力を行ったり、モデムを介してファクシミリ送信を行うのもこの部分とする。

【0027】106は音声入力部で、マイクロホン等を使用し、音声の入力を行い、デジタルデータとして音声認識部に送る。

【0028】107は音声認識部で、音声入力部106からの音声データを入力し、辞書データへの登録や音声のパターンマッチング、コマンドへの変換等を行う。変換されたコマンドは、操作部101からのコマンドと同様に、ドライバ部102へと送られる。

【0029】次に、本発明の特徴である音声認識部107についての詳細な説明を、その内部処理構成ブロック図を示す図2を用いて行う(本図も図1と共に、後述する第2～第5実施例にも共用するものとする)。

【0030】201は音声データ入力部で、この音声入力部201からのデジタル音声データを入力する。

【0031】202は操作部情報入力部で、操作部101から入力されたコマンド情報の入力を行う。

【0032】203は辞書データ蓄積部で、操作部情報と音声データとを対応させて、メモリやハードディスク等の記憶媒体に記憶しておく。

【0033】204はパターンマッチング部で、音声データ入力部201からのデータと、辞書データ蓄積部203からの音声辞書データとのパターンマッチングを行い、各コマンドに対する類似度を出力する。

【0034】例えば、入力された音声データが、「いちまい」であったとき、音声辞書データの「いちまい」との類似度は100、「はちまい」との類似度は80、「さんまい」との類似度は30、「コピー」との類似度は5、などと出力を行う。

【0035】205はドライバ情報入力部で、辞書情報の登録や通常コピー動作等の情報をドライバ部102から入力を行う。

【0036】206はコマンド処理部で、パターンマッチング部204からの情報とコマンド情報入力部205の情報とを受けて、通常コピーの場合は、入力された音声コマンドがいずれのコマンドに相当するかを判断し、

コマンド出力部 207 に送る。また、辞書情報の登録時には、類似度の高いコマンドのいくつかをコマンド出力部 207 に送る等の処理を行う。

【0037】207 はコマンド出力部で、コマンド処理部 206 からのコマンド情報を、通常コピーの場合はドライバ部 102 や操作部 101 のモード表示部分へ送る。また、辞書情報の登録時には、類似したコマンドや警告音を発生する信号を操作部 101 へ送る。

【0038】次に具体的な例を挙げて処理内容を説明する：音声コマンドの登録を行う際には、実際に所定のキーを押してから発声を行う。この時、音声データ入力部 201 からは、その時の音声データが、また、操作部情報入力部 202 からは、その時押されているキーの情報が入力される。

【0039】例えば、「白黒コピーを 1 枚」行う時には、図 8 の「コピー」キーを押しながら、「コピー」と発声する。すると、操作部 101 からドライバ部 102 に「コピー」というコマンドが伝えられると共に、音声認識部 107 の操作部情報入力部 202 にも同様にコマンドが伝えられる。

【0040】また、音声入力部 106 からの音声データは、音声データ入力部 201 に入力され、コマンド情報と共に辞書データ蓄積部 203 に入力される。辞書データ蓄積部 203 では、「コピー」というコマンドに対応して、入力された音声データを記憶しておく。

【0041】以下、同様に「白黒」というコマンドと音声データ、「1 枚」というコマンドと音声データを登録すると同時に設定を行っていく。

【0042】スタートキーを押しながら、「スタート」と発声することによって、「スタート」という音声コマンドが登録されると共に、コピー出力を行うことができる。

【0043】以上の作業により、実際にコピーを行いながら「コピー」、「白黒」、「1 枚」、「スタート」といった音声コマンドが登録される。

【0044】以上の音声コマンド登録時の動作シーケンスフローチャートは、図 3 のシーケンス実線矢印部分に該当する。音声入力（ステップ S301）を、図 2 の音声データ入力部 201 で行い、その時の操作部の状況を操作部状況入力部 202 で検知し、キー入力と同時に

行われているか否かの判断を行う（ステップ S302）。【0045】キー入力と同時に行われている場合、音声データ登録モード（ステップ S303）となる。パターンマッチング部 204 によって音声データのパターンマッチングを行い（ステップ S304）、類似コマンドが無かった場合（ステップ S305）には、入力された音声データをコマンドとして辞書データ蓄積部 203 に登録し（ステップ S306）、そのコマンドを実行（ステップ S307）する。

【0046】なお、以上の各ステップにおける NO を示

すシーケンス点線矢印で示す各ステップの逐次詳細説明は、図示の通りであり逐次詳細の重複説明は省略する。

【0047】以上の方法によって登録されたコマンドは、キー操作を行わなくても音声のみで使うことができることとなる。

【0048】つまり、白黒コピーを 1 枚採りたい場合、「コピー」と発声すれば「コピー」モードとなり、「白黒」と発声することによって「白黒」モードになる。その後、「1 枚」、「スタート」と発声することによって、白黒コピーの動作を行うことができる。

【0049】この時の音声コマンド実行時の動作シーケンスフローチャートは、図 4 のシーケンス実線矢印部分に該当する。音声入力（ステップ S401）を図 2 の音声データ入力部 201 で行い、その時の操作部 101 の状況を操作部状況入力部 202 で検知し、キー入力が行われているか否かの判断を行う（ステップ S402）。

【0050】キー入力が行われていない場合、音声データ実行モード（ステップ S403）となる。パターンマッチング部 204 によって音声データのパターンマッチングを行い（ステップ S404）、登録されているコマンドに類似したものが 1 つだけあった場合（ステップ S405・S406）には、入力された音声データをコマンドとして実行（ステップ S407）する。

【0051】このときの音声コマンドの認識について、詳細に説明する：音声コマンドの認識は、図 2 の音声データ入力部 201 からの音声データと、辞書データ蓄積部 203 に格納してある音声データとのパターンマッチングを、パターンマッチング部 204 によって行い、各コマンドとの類似度を算出しコマンド処理部 206 に出力する。この場合、登録してある辞書データと完全に一致していた場合、類似度を 100 とする。

【0052】例えば、「白黒」といった音声データ入力があった場合、辞書データ蓄積部 203 にある音声辞書データの「白黒」とは一致度が 95、「コピー」との一致度は 5、「1 枚」との一致度は 10、「スタート」との一致度は 5、などといった情報がコマンド処理部 206 に送られる。

【0053】コマンド処理部 206 では、最も一致度の高いものを音声入力されたコマンドと判断して、コマンド出力部 207 に送る。

【0054】コマンド出力部 207 では、そのコマンドをドライバ部 102 に送ることによって、本体の設定を行うこととなる。

【0055】上記 2 つの動作フロー図 3、4 の組み合わせにより、次回カラーコピーを 1 枚とる場合、「コピー」と発声すれば「コピー」モードとなり、「カラー」と発音しながら「カラー」キーを押すことによって「カラー」モードになると共に、新たに「カラー」という音声コマンドを登録する。

【0056】その後「1枚」、「スタート」と発声することによって、「カラー」という発声コマンドを新たに登録すると共に、カラーコピーの動作をも行うことができる。

【0057】以上のように音声コマンドのデータ入力を行うことによって、必要なコマンドを、必要な時に登録することができ、普段は使用しないコマンドの登録を防止することによって、辞書データ登録部203に使用する記憶媒体の容量増大を抑制することができる。

【0058】なお、本実施例の音声コマンド実行時の図4各ステップにおける各シーケンス点線矢印で示す各ステップの逐次詳細説明は、図示の通りであり、逐次詳細の重複説明は省略する。

【0059】（第2の実施例）本発明に係る第2の実施例を、それぞれ図1、図2を共用して説明する：もし、音声コマンドとして登録されていないコマンドが入力された時の処理は、以下のようになる。

【0060】まず、音声コマンドは図1、2における操作部101のキー図8が押されない状態で入力されることによって、新たな音声コマンドの登録ではないと判断

【0061】次に、図2におけるパターンマッチング部204によって登録されている音声辞書データとのパターンマッチングを行い、各コマンドとの類似度を算出しコマンド処理部206に送る。

【0062】この場合、登録されているコマンドとは異なるため、類似度は全て低い値となる。完全に一致している場合、類似度を100とした場合に、80以上をコマンドの候補、80未満は登録されていないコマンドの入力と判断すると、全てのマッチングを行った結果、登録されているコマンドの候補は無い状態となる。

【0063】そのような場合、コマンド処理部によって入力された音声データは、登録されていないものと判断し、コマンド出力部207から、操作部101の液晶表示部に「未登録です。登録するキーを押して下さい。」の表示を行うことによって、使用者に報知するようにドライバ部102に情報を送る。

【0064】使用者は、未登録であることを確認した後、操作部101のキーを押すことにより新たにコマンドを登録することとなり、また同時にそのコマンドは実行され、次の処理に進む。

【0065】以上の本第2実施例の音声登録時の動作シーケンスフローチャートは、図5のシーケンス実線矢印部分に該当する。音声入力（ステップS501）を図2の音声データ入力部201で行い、その時の操作部101の状況を操作部状況入力部202で検知し、キー入力と同時に進行しているか否かの判断を行う（ステップS502）。

【0066】キー入力と同時に進行していない場合、音声データ実行モード（ステップS503）となる。パタ

ーンマッチング部204によって音声データのパターンマッチングを行い（ステップS504）、類似コマンドが無かった場合（ステップS506）には、コマンドが未登録であることを使用者にメッセージとして操作部101から表示し（ステップS506）、その後キー入力のあったキーを音声コマンドとして登録する（ステップS507）。それと共にコマンドを実行する（ステップS508）。

【0067】以上のように、音声コマンドのデータ入力を行うことによって、音声コマンドとして登録されていないコマンドが音声入力されたときに簡単に新たなコマンドとして登録することができる。

【0068】本第2実施例の音声コマンド登録時図5の各ステップにおける各シーケンス点線矢印で示す各ステップの逐次詳細説明は図示の通りであり、逐次詳細の重複説明は省略する。

【0069】（第3の実施例）次に本発明の第3の実施例をそれぞれ前記図1、図2を共用して説明する：音声コマンドの登録を行う場合、既に登録されているコマンドを更に登録しようとしてしまった時の動作について説明する。

【0070】まず、音声コマンドが操作部101のキーが押された状態で入力されることによって、新たな音声コマンドの登録であると判断する。

【0071】次に、パターンマッチング部204によって登録されている音声辞書データとのパターンマッチングを行い、各コマンドとの類似度を算出し、コマンド処理部206に送る。

【0072】この場合、既に登録されているコマンドである場合、類似度の高い値となるコマンドが存在することとなる。

【0073】そのような場合、コマンド出力部207から、操作部101の液晶表示部（図8）に、類似度の高かったコマンドを表示することによって、使用者に確認を促すように、ドライバ部102に情報を送る。

【0074】具体例を挙げると、「1枚」というコマンドが既に登録されているにもかかわらず、再び登録しようとした場合、使用者は操作部101のキー（図8）を押しながら「1枚」と発声する。

【0075】前記第1の実施例において説明した方法と同様に、音声認識部107によってパターンマッチングを行い、類似度の算出を行うと、「1枚」というコマンドの類似度が90となり、80以上をコマンドの候補、80未満は登録されていないコマンドの入力と判断すると、それはすでに登録されているコマンドと判断される。

【0076】このような場合、図9にその一例を示すように液晶タッチパネル（図8）の表示部に「今のコマンドは「1枚」ですか？」という表示と「YES」、「NO」の選択を行う表示とを行い、使用者に登録済みのデ

ータと類似していることを報知する。

【0077】「YES」のキーを押すことによって、使用者は、既に登録されていることが確認できると共に、ドライバ部102は、入力されたコマンドを受け付けて、その後の処理を行うこととする。「NO」の場合は再登録となる。

【0078】以上の動作シーケンスフローチャートは、図6のシーケンス実線矢印部分に該当する。音声入力（ステップS601）を音声データ入力部201で行い、その時の操作部101の状況を操作部情報入力部202で検知し、キー入力が行われているか否かの判断を行う（ステップS602）。

【0079】キー入力が行われている場合、音声データ登録モード（ステップS603）となる。パターンマッチング部204によって音声データのパターンマッチングを行い（ステップS604）、類似コマンドがあった場合（ステップS605）には、コマンドが登録済みであることを使用者にメッセージとして操作部101から示し（ステップS606）、その後「YES」のキー入力があったら（ステップS607）、コマンドを実行する（ステップS608）。

【0080】以上のように、音声コマンドのデータ処理を行うことによって、すでに登録されているコマンドを更に登録しようとしてしまった時に、2重に登録してしまうことを防止することができる。

【0081】なお、本第3実施例の音声コマンド登録時の図6各ステップにおけるシーケンス点線矢印で示す各ステップの逐次詳細説明は、図示の通りであり、逐次詳細の重複説明は省略する。

【0082】（第4の実施例）次に本発明に係る第4の実施例を、それぞれ図1、図2を共用して説明する：未登録のコマンドを登録しようとした場合、既に登録されているコマンドと類似している音声データが入力された時の動作について説明する。

【0083】まず、音声コマンドが操作部101のキーが押された状態で入力されることによって、新たな音声コマンドの登録であると判断する。

【0084】次に、パターンマッチング部204によって登録されている音声辞書データとのパターンマッチングを行い、各コマンドとの類似度を算出し、コマンド処理部206に送る。

【0085】この場合、既に登録されているコマンドで、類似度の高い値となるコマンドが存在した場合、コマンド出力部207から、操作部101の液晶表示部（図8）に、類似度の高かったコマンドを表示することによって使用者に確認を促すように、ドライバ部102に情報を送る。

【0086】具体例を挙げると、「8枚」というコマンドを新たに登録しようとした場合、使用者は操作部101のキー（図8）を押しながら「8枚」と発音する。し

かしながら、音声の仕方により、すでに登録済みの「1枚」という音声コマンド共、類似度が大きくなったとする。

【0087】前記第1の実施例において説明した方法と同様に、音声認識部107によってパターンマッチングを行い、類似度の算出を行うと、「1枚」というコマンドの類似度が85となり、80以上をコマンドの候補、80未満は登録されていないコマンドの入力と判断すると、それは既に登録されているコマンドと同じレベルの類似度と判断される。

【0088】この場合、その後音声によるコマンド入力を行おうとした際に、誤認識し易くなる。

【0089】このような場合、図10にその一例を示すように、液晶タッチパネル（図8）の表示部に「今のコマンドは「1枚」と類似しています。登録し直しますか？」という表示と「YES」、「NO」の選択を行う表示とを行い、使用者に登録済みのデータと類似していることを報知する。

【0090】「YES」のキーを選択すると、再登録を行うこととなるし、「NO」のキーを押すと、そのまま音声コマンドの登録を行うと共に、ドライバ部102は、入力されたコマンドを受け付けて、その後の処理を行うこととする。

【0091】以上の動作シーケンスのフローチャートは、前記第3実施例の図6の実線部分と同様にできる。音声入力（ステップS601）を音声データ入力部201で行い、その時の操作部101の状況を操作部状況入力部202で検知し、キー入力が行われているか否かの判断を行う（ステップS602）。

【0092】キー入力が行われている場合、音声データ登録モード（ステップS603）となる。パターンマッチング部204によって音声データのパターンマッチングを行い（ステップS604）、類似コマンドがあった場合（ステップS605）には、類似したコマンドが登録済みであることを使用者にメッセージとして操作部101から示し（ステップS606）、その後キー入力のあったキーを音声データと対応したコマンドとして登録し（ステップS607）、コマンドを実行する（ステップS608）。

【0093】以上のように、音声コマンドのデータ処理を行うことによって、新たに登録しようとした音声データが既に登録されているコマンドと類似していた時にも、正確に登録を行うことができる。

【0094】なお、本第4実施例においても、各シーケンスの点線矢印で示す各ステップの詳細は、前記第3実施例における同様に、詳細重複説明は省略する。

【0095】（第5の実施例）さらに、本発明に係る第5の実施例を、それぞれ図1、2を共用して説明する：音声コマンドの認識を行う場合、音声入力されたコマンドに対して、複数の類似したコマンドの候補が出てしま

い、使用者の意志に反した動作をしてしまうことがある。そのような誤動作を防ぐために、以下のような処理を行う。

【0096】まず、音声コマンドは、操作部101(図2、図8)のキーが押されない状態で入力されることによって、音声コマンドによる設定であると判断する。

【0097】次に、パターンマッチング部204(図2)によって登録されている音声辞書データとのパターンマッチングを行い、各コマンドとの類似度を算出し、コマンド処理部206に送る。

【0098】この場合、実際に設定したいコマンドと類似した音声入力となされた場合、類似度の高い値となるコマンドが複数存在することとなる。

【0099】そのような場合、コマンド出力部207から、操作部101の液晶表示部に、類似度の高かった複数のコマンドを表示することによって、使用者に選択を促すようにドライバ部102に情報を送る。

【0100】具体例を挙げると、「一枚」というコマンドを設定したいと思い、使用者は「1枚」と発音する。

【0101】しかしながら、登録時との発音の違い等により、仮に、「8枚」という音声コマンドとも類似度が大きくなったとする。

【0102】前記第1の実施例において説明した方法と同様に、音声認識部107によってパターンマッチングを行い、類似度の算出を行うと、「1枚」と云うコマンドの類似度が90となり、「8枚」と云うコマンドの類似度が85となったとする。類似度が80以上をコマンドの候補、80未満は登録されていないコマンドの入力と判断すると、両方共コマンドの候補として挙げられると判断される。

【0103】このような場合、図11にその一例を示すように、液晶タッチパネルの表示部に「今のコマンドはどちらですか? 1:「1枚」・2:「8枚」と表示を行い、使用者にどちらのコマンドか選択して貰う。それによって選択された設定をドライバ部102は、入力されたコマンドとして、その後の処理を行うこととする。

【0104】以上の動作シーケンスフローチャートは、図7動作シーケンス実線矢印部分のようになる。(前記第3、4実施例の図6の破線矢印部分の一部に相当する)。音声入力(ステップS701)を音声データ入力部201(図2)で行い、その時の操作部101の状況を操作部状況入力部202で検知し、キー入力と同時に進行されているか否かの判断を行う(ステップS702)。キー入力が行われていない場合、音声データ実行モード(ステップS703)となる。パターンマッチング部204によって音声データのパターンマッチングを行い(ステップS704)、類似コマンドが複数あった場合、(ステップS705・S706)には、類似した複数のコマンドが登録済みであることを使用者にメッセージとして操作部101から示し(ステップS70

7)、その後キー入力によって指定されたコマンド(ステップS708)を実行コマンドとして実行する(ステップS709)。

【0105】以上のように、音声コマンドのデータ処理を行うことによって、入力された音声コマンドとの類似度が高いものが複数登録された場合も、正確な設定ができることとなる。

【0106】

【発明の効果】以上説明したように、本発明によれば、前記音声によるコマンドの認識を行い動作する画像処理装置において、辞書データ蓄積手段にデータを登録する際に、実際に操作し乍ら音声データの登録を行うことが必要なコマンドを、必要な時に登録することができ、普段は使用しないコマンドの登録を防止することによって、辞書データ蓄積手段に使用する記憶媒体の容量を抑えることが可能となる。

【0107】また音声データを登録する際に、前記辞書データ蓄積手段のデータに、既に登録されているデータと同じデータを登録しようとした場合、そのことを使用者に表示することによって、2重のデータ登録を防ぐことが可能となる。

【0108】また、音声によって設定を行う際に、未登録のコマンドを音声によって設定しようとした場合、未登録であることを使用者に表示し、その都度、登録を行っていくことが可能となる。

【0109】さらにまた、音声によって設定を行う際に、音声入力されたコマンドが、複数のコマンドと類似していた場合、それらのコマンドを表示し、使用者に選択を促すことにより、正確なコマンド入力が可能となる。

【図面の簡単な説明】

【図1】 各実施例の複写機の構成ブロック図

【図2】 図1の音声認識部の内部処理構成ブロック図

【図3】 第1実施例の音声コマンド登録時の動作シーケンスフローチャート

【図4】 第1実施例の音声コマンド実行時の動作シーケンスフローチャートを示す図である

【図5】 第2実施例の音声コマンド登録時の動作シーケンスフローチャート

【図6】 第3、第4実施例の音声コマンド登録時の動作シーケンスフローチャート

【図7】 第5実施例の音声コマンド登録時の動作シーケンスフローチャート

【図8】 図1複写機の操作部平面図例

【図9】 第3実施例の操作部の液晶タッチパネルの表示例

【図10】 第4の実施例の操作部の液晶タッチパネルの表示例

【図11】 第5の実施例の操作部の液晶タッチパネルの表示例



## 【符号の説明】

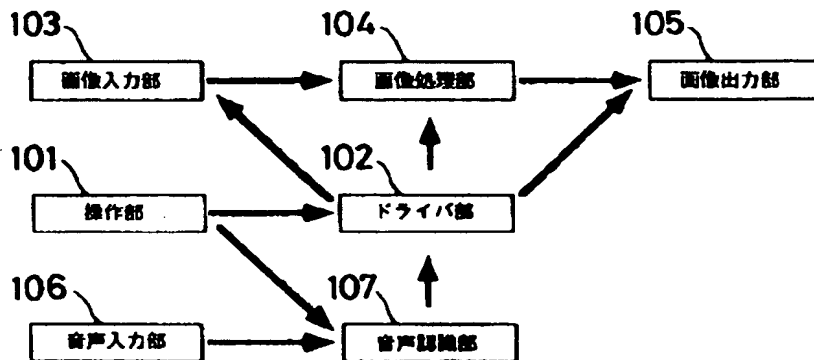
101 操作部  
102 ドライバ部  
103 画像入力部  
104 画像処理部  
105 画像出力部  
106 音声入力部  
107 音声認識部

201 音声データ入力部  
202 操作部情報入力部  
203 辞書データ蓄積部  
204 パターンマッチング部  
205 ドライバ情報入力部  
206 コマンド処理部  
207 コマンド出力部

【図1】

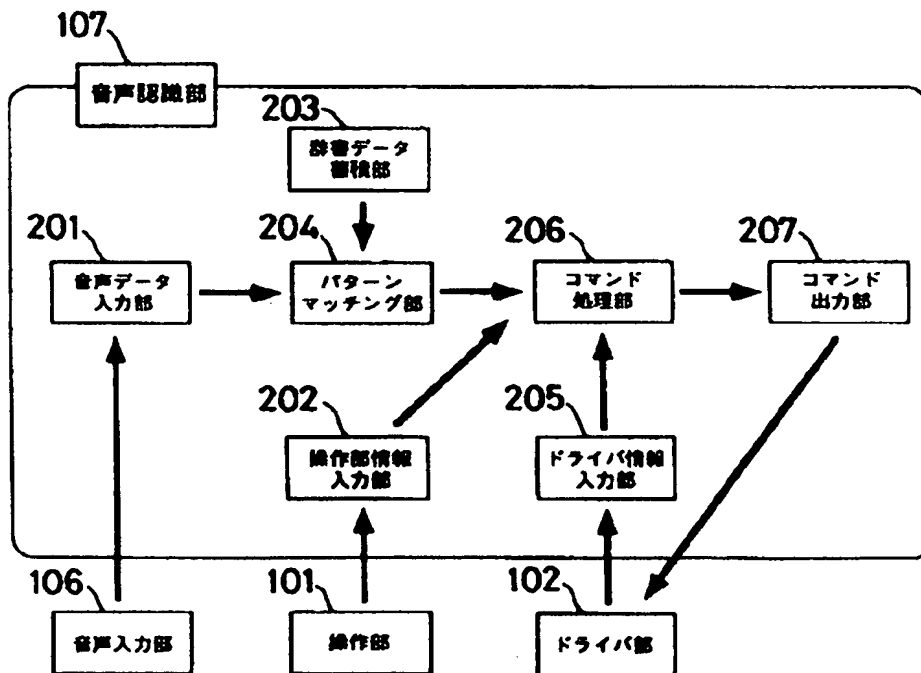
【図9】

各実施例の複写機の概要構成ブロック図

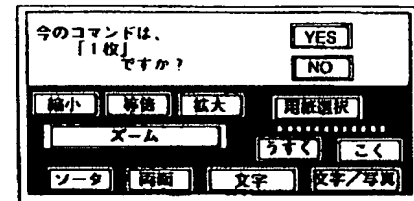


【図2】

図1の音声認識部の内部処理構成ブロック図

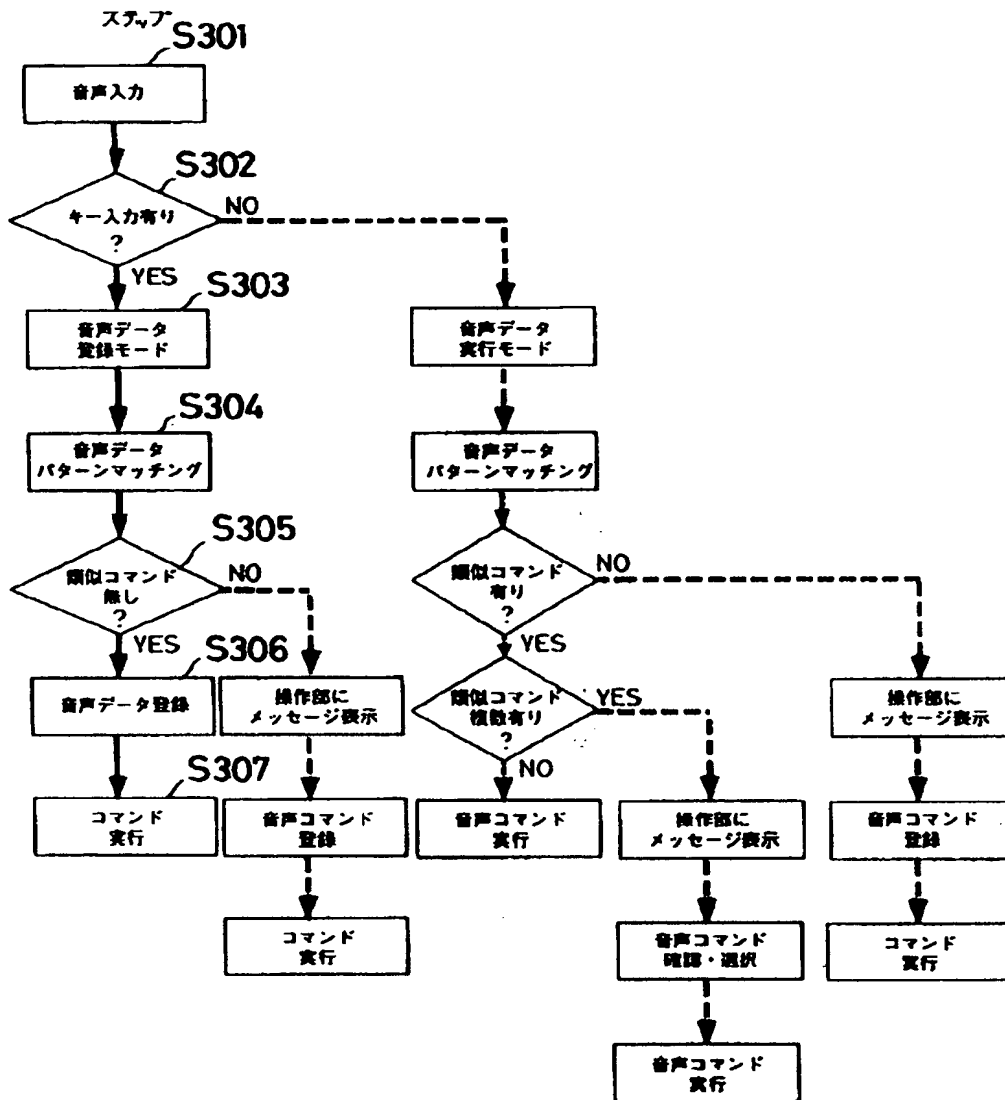


第3実施例の操作部の液晶タッチパネルの表示例



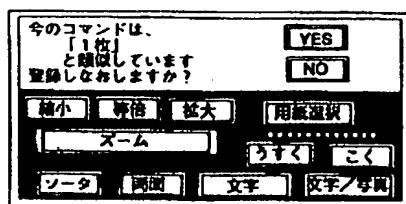
【図3】

## 第1実施例の音声コマンド登録時の動作シーケンスフローチャート



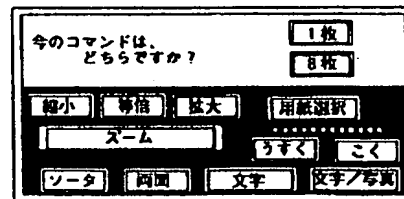
【図10】

## 第4実施例の操作部の液晶タッチパネルの表示例



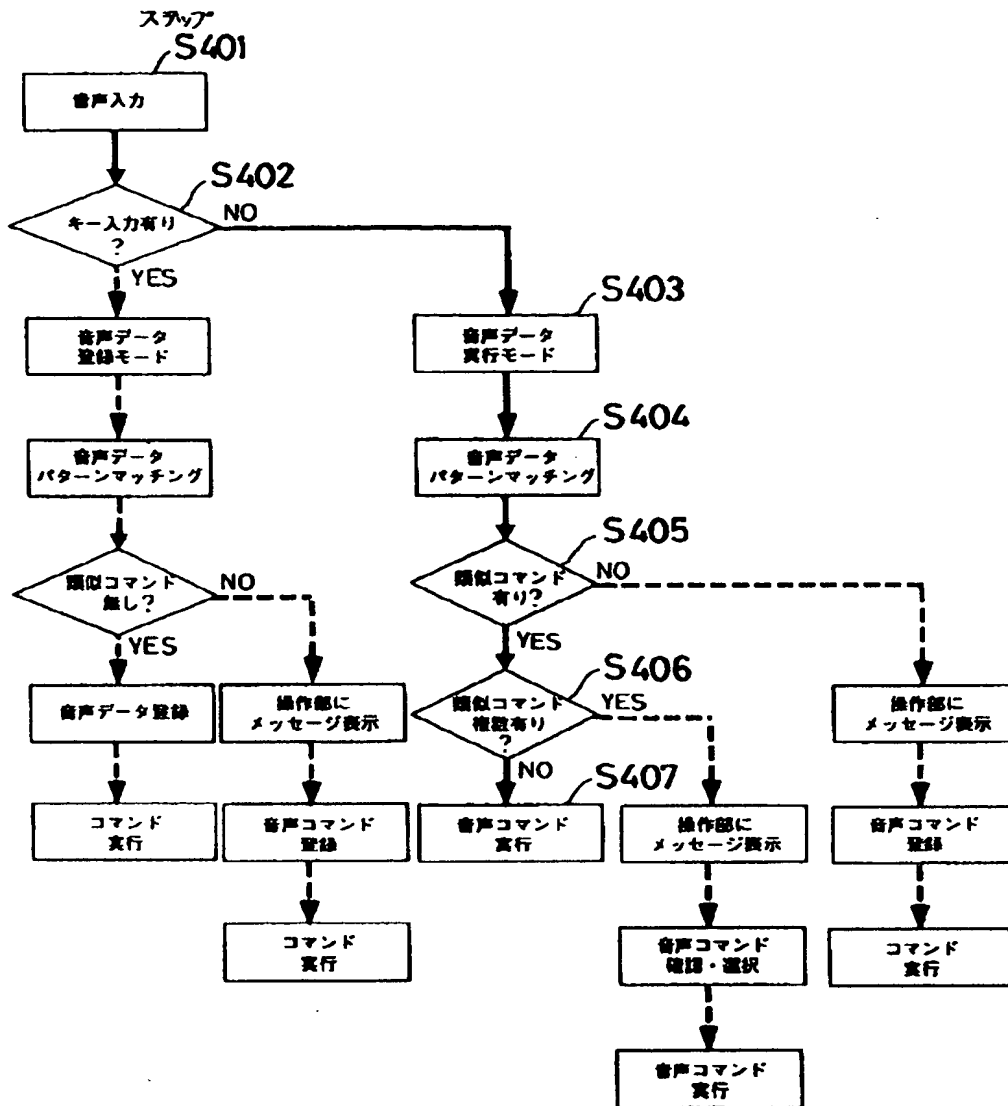
【図11】

## 第5実施例の操作部の液晶タッチパネルの表示例



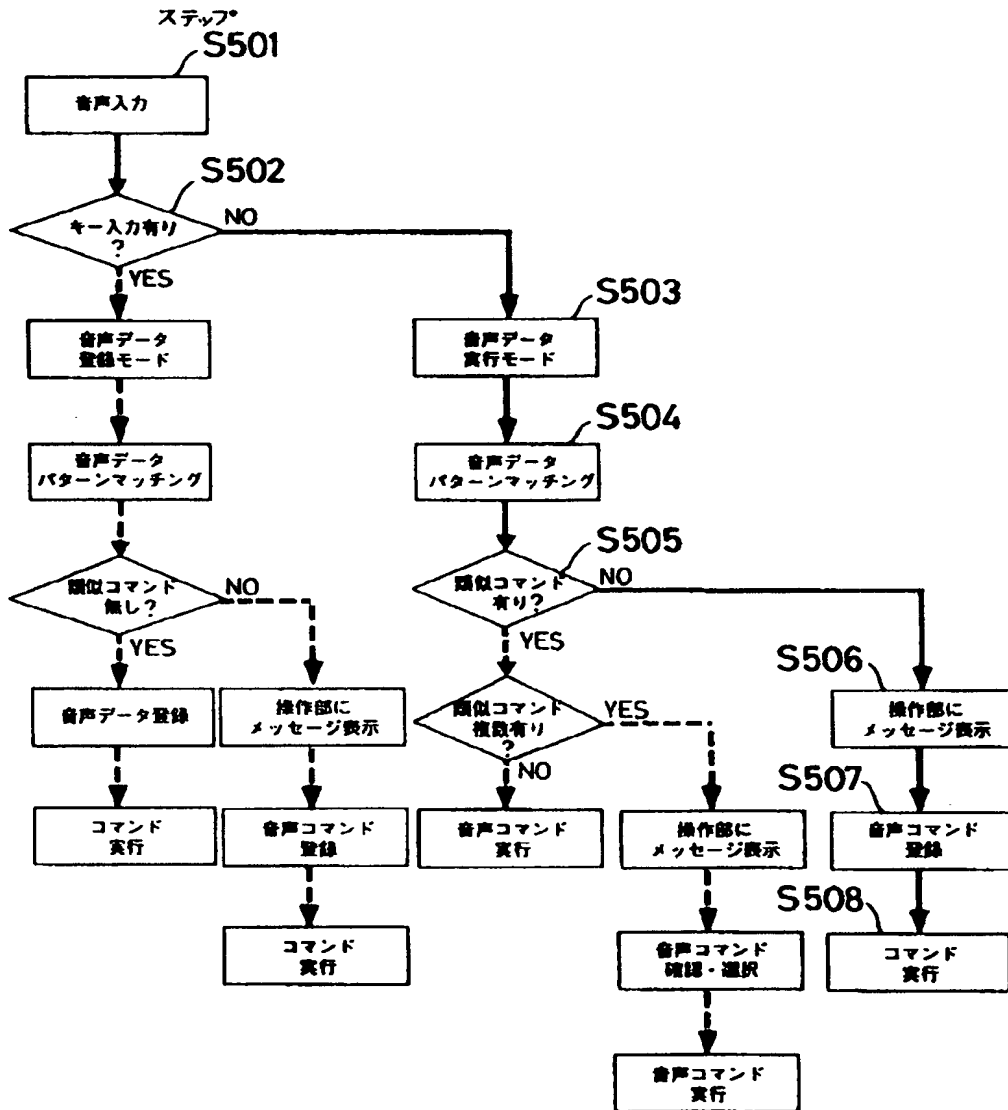
【図4】

## 第1実施例の音声コマンド実行時の動作シーケンスフローチャート



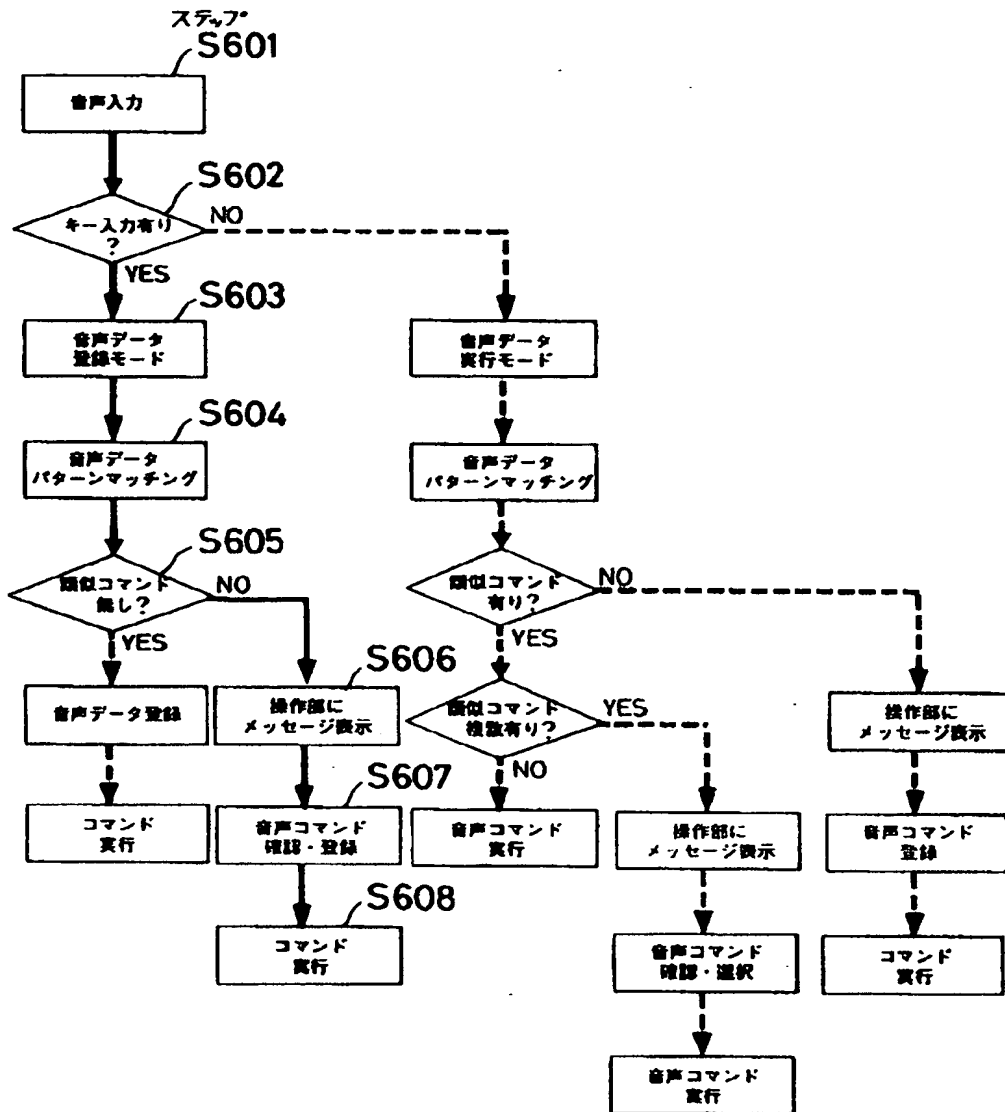
【図5】

## 第2実施例の音声コマンド登録時の動作シーケンスフローチャート



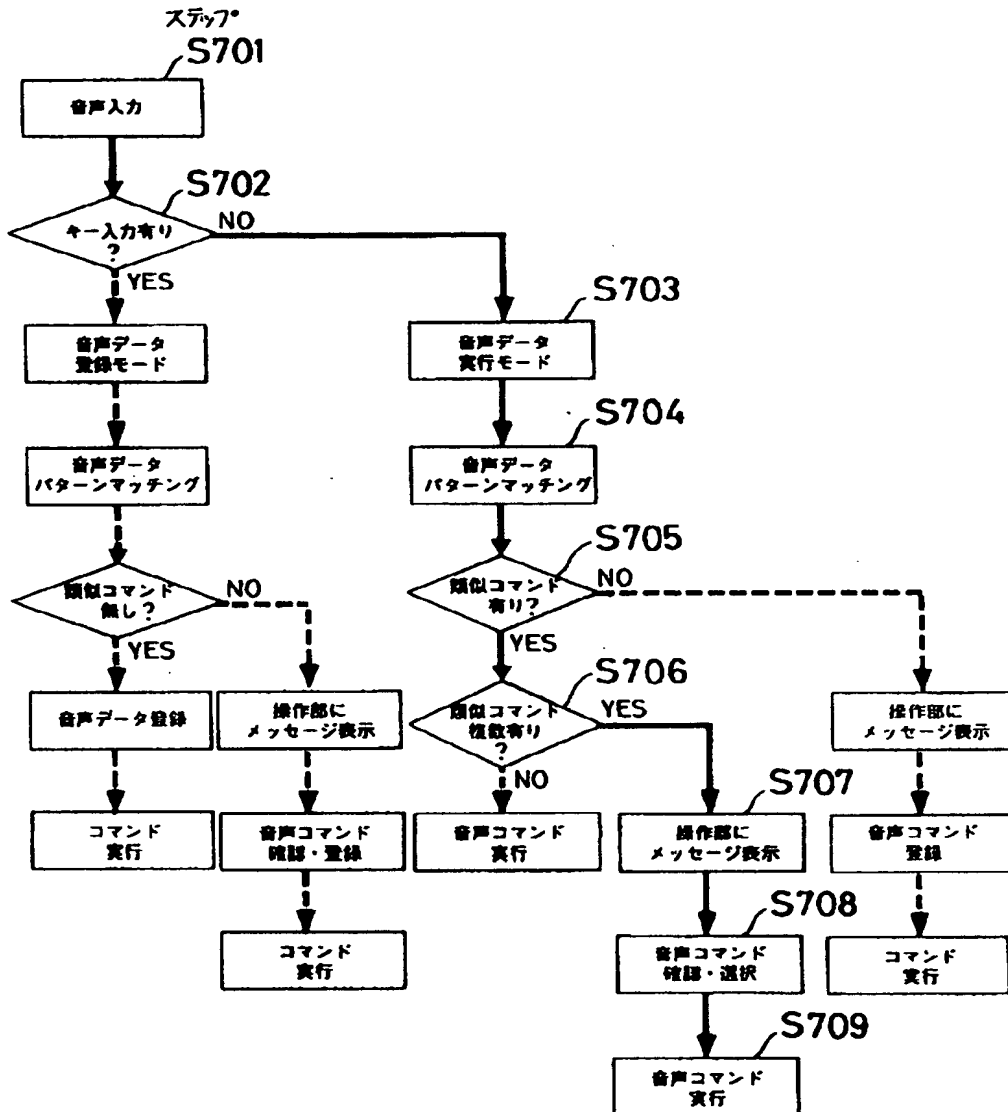
【図6】

## 第3, 第4 実施例の音声コマンド登録時の動作シーケンスフローチャート



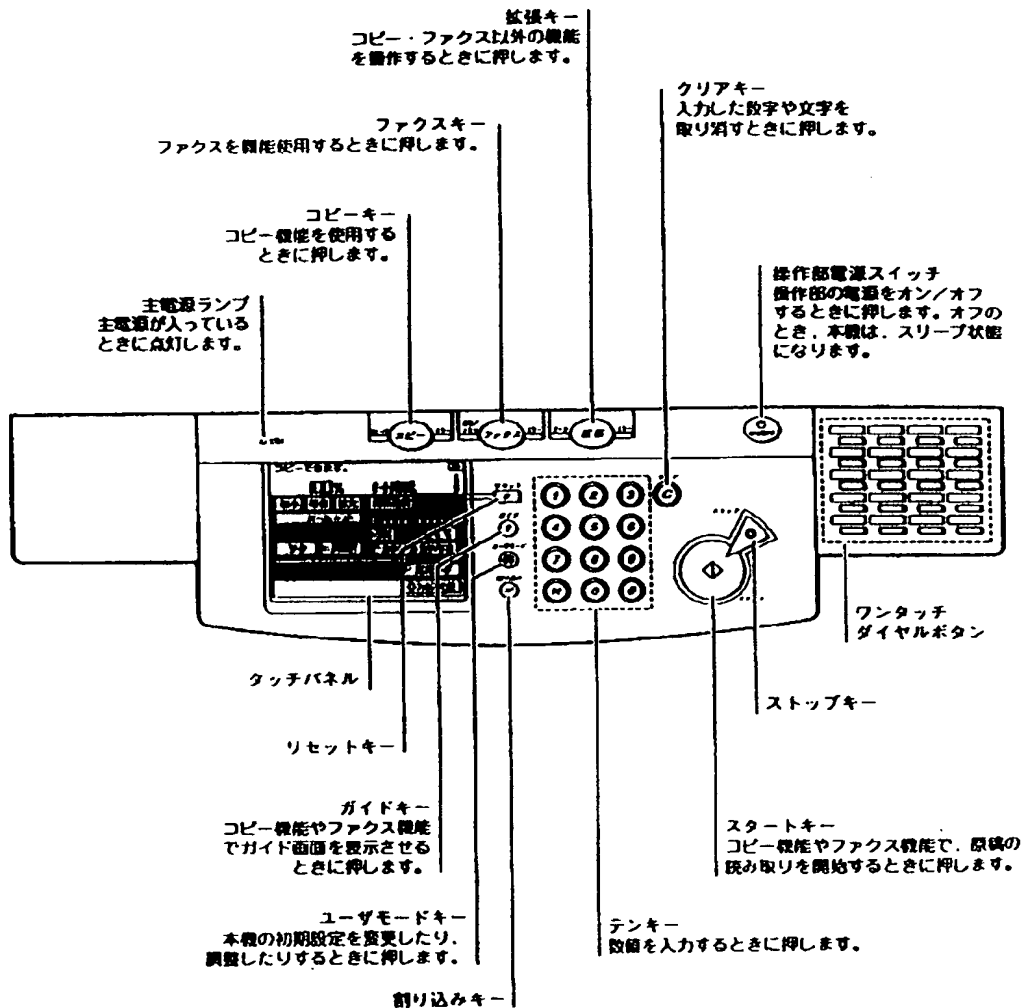
【図 7】

## 第5実施例の音声コマンド登録時の動作シーケンスフローチャート



【図 8】

図 1 複写機の操作部平面図例



フロントページの続き

(51) Int. Cl.<sup>7</sup>

識別記号

F I

テマコード (参考)

H 0 4 N 1/00

G 0 6 F 3/16

3 2 0 B

// G 0 6 F 3/16

3 2 0

B 4 1 J 29/00

J

Fターム(参考) 2C061 AP01 AP03 AP04 AP07 AQ05  
AQ06 CQ02 CQ04 CQ24 CQ25  
2H027 GA19 GA20 GA25 GA26 ZA07  
5B021 AA05 AA19 CC05 PP04 PP06  
QQ02 QQ04  
5C062 AA02 AA05 AB02 AB11 AB20  
AB25 AB26 AB41 AB42 AC02  
AC03 AC05 AC21 AC60 AE11  
AF07 AF15 BA00  
5D015 GG04 KK01 LL00



# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-029585

(43)Date of publication of application : 28.01.2000

(51)Int.Cl. G06F 3/00  
B41J 29/17  
G03G 21/00  
G06F 3/12  
G10L 15/00  
H04N 1/00  
// G06F 3/16

(21)Application number : 10-193069

(71)Applicant : CANON INC

(22)Date of filing : 08.07.1998

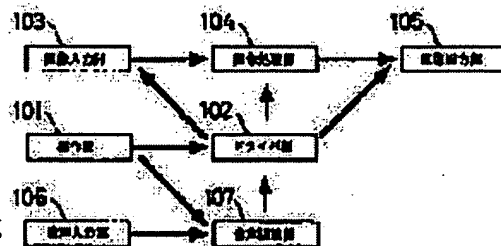
(72)Inventor : AOYANAGI TAKESHI

## (54) VOICE COMMAND RECOGNIZING IMAGE PROCESSOR

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a means for relaxing excess in the capacity of a storage medium such as a memory for previously registered audio dictionary data in the case of using a copy machine or a combined machine consisting of facsimile and printer to be operated while recognizing an audio input command from many and unspecified persons.

SOLUTION: When registering reference data in the dictionary data, these data are registered while actually pressing the button of an operating part 101 or the like and when next using the same command, this device can be operated just by voice. In the case of performing registration, pattern machining is executed as well and when the same command as a registered command, any similar command or similar voice is inputted an alarm is issued to the user and in the case of dictionary data having high similarity, these data are more exactly recognized by any other judging element such as a fingerprint or palm.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or

BEST AVAILABLE COPY

application converted registration]  
[Date of final disposal for application]  
[Patent number]  
[Date of registration]  
[Number of appeal against examiner's decision  
of rejection]  
[Date of requesting appeal against examiner's  
decision of rejection]  
[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office